Radiology

Automated Aneurysm Detection: Emerging from the Shallow End of the Deep Learning Pool

David F. Kallmes, MD • Bradley J. Erickson, MD, PhD

Dr Kallmes is professor of radiology and neurosurgery at Mayo Clinic and served as deputy editor for the neuroradiology section at *Radiology* for 8 years. As an interventional neuroradiologist, he has been involved in both aneurysm research and patient care for 25 years. He supervises an NIH-funded preclinical laboratory focused on aneurysm pathophysiology and treatment. He was awarded the Outstanding Contributions in Research Award at the 2019 ASNR Annual Meeting.

Dr Erickson is professor of radiology and informatics at the Mayo Clinic. He directs the radiology AI laboratory and has been a clinical neuroradiologist for 25 years.

Missing a brain aneurysm at CT angiography or MR angiography represents a source of unending anxiety among the neuroimaging community, an anxiety exacerbated by the fact that accuracy of these tests fast diminishes with decreasing aneurysm size. The proliferation of CT angiography for ever-expanding indications adds to the anxiety for general radiologists (one could argue even more than for neuroradiologists). Automated detection of aneurysms by using computer-aided diagnosis (CAD) was initially studied in 2004 (1), but such techniques failed to achieve clinical impact given the lack of user-friendly interfaces, high rates of false-positive diagnoses, and low sensitivity for small aneurysms. The need to scrutinize numerous "hits" from the algorithm, without substantial improvement in sensitivity, may result in reader fatigue that defeats the purpose of the CAD. In recent years, multiple groups have evaluated deep learning, often with convolutional neural networks, in aneurysm detection (2). But these more recent studies have generated criticism for their small numbers of aneurysms (3), especially ruptured aneurysms (4), dearth of negative cases (5), limited cross-vendor and multi-institutional data sets (6), and lack of reference standard digital subtraction angiography (4). Further, these latter studies still suffer from high rates of false-positive diagnoses (5).

Against this backdrop of years (now near decades) of unrealized hopes of rapid, reliable, and easy aneurysm detection, in this issue of *Radiology*, Yang et al (7) report their experience with a convolutional neural network approach. Compared with previous literature, their study is enhanced by large data sets. They include reasonable numbers of small aneurysms (mean size of 5 mm for the training set and 4 mm for the validation set), external validation (a multi-institutional data set distinct from the training and validation sets but still the same institutions used for training), and the use of general, rather than subspecialty trained, radiologists for model validation. Their algorithm had a high sensitivity of 97.5%, with most difficulty in the setting of less than 3-mm aneurysms near the bony structures of the skull base. The overall area under the receiver operating characteristic (ROC) curve for the external validation readings were better with the algorithm than without it, although two of the four readers achieved a 5% improvement in ROC curve area whereas the other two readers achieved no improvement in sensitivity. Also, durations of the readings were slightly shorter with the algorithm than without it, ranging from 8–18 seconds shorter with the algorithm for the internal validation cohort. Of note, almost the entirety of benefit in sensitivity with the algorithm was in internal carotid artery aneurysms, which is expected given the proximity to bony structures and potential venous contamination, and those of less than 7-mm diameter, with between three and eight additional detected aneurysms among 85 such aneurysms in the entire cohort. Such aneurysms are and continue to be the bane of aneurysm detection with any method short of digital subtraction angiography. The authors conclude: "The proposed deep learning algorithm assisted radiologists in detecting cerebral aneurysms on CT angiography images, resulting in a higher detection rate…" than without the algorithm.

Do these conclusions indicate that we have (finally) entered the era of artificial intelligence–assisted aneurysm detection? Probably not yet. At best, this proposed algorithm is an interpretation aid rather than a diagnostic aid because the authors used no true standard of reference (ie, conventional angiography). Another headwind against the algorithm was the lack of so-called negative cases in the training and initial validation sets, such that the machine would elevate an equivocal finding to a positive one, exaggerating the sensitivity. Most concerning is the ongoing high rate of false-positive findings, with 13.8 per case to reach the sensitivity of 97.5%. (A more manageable

From the Department of Radiology, Mayo Clinic, 200 First St SW, Rochester, MN 55905. Received September 24, 2020; revision requested October 5; revision received October 13; accepted October 19. **Address correspondence to** D.F.K. (e-mail: *kallmes.david@mayo.edu*).

Conflicts of interest are listed at the end of this article.

See also the article by Yang et al in this issue.

Radiology 2021; 298:164–165 • https://doi.org/10.1148/radiol.2020203853 • Content codes: IN CT •

number of false-positive findings of four to five per case dropped sensitivities to the low 90% range, not a major value added over the unaided reader.) So little time difference with and without the algorithm seems quite surprising, especially given the high number of false-positive diagnoses per case (one reader read the cases in an average of 30 seconds both with and without this tool). Given that these studies are usually around 100 source images, not considering reformats, that is one-third of a second per image. That makes one wonder whether the CAD tool was even considered by that reader. Even with the other viewers, the time benefit was only 3.6 seconds, and the performance benefit was neither of clinical nor statistical significance.

Important clinical features of cerebral aneurysms also require further explanation. First, not all cerebral aneurysms are the same. Size, location, family history, smoking history, and especially rupture status represent essential features that guide care. Indeed, a ruptured aneurysm is a medical emergency, with rerupture rates of 20% by 14 days and 70% mortality with such rehemorrhage (8). Missing even a small ruptured aneurysm is unacceptable. Conversely, a small, anterior circulation unruptured aneurysm (ie, the same type of aneurysm most impacted by the algorithm) in a patient without a personal or family history of aneurysm rupture carries an annual rate of rupture of 0.07% (9). In the study by Yang et al, it is unclear how the algorithm performed in ruptured versus unruptured aneurysms. But common sense—and local standard of care—should mandate definitive digital subtraction angiography in almost any case of spontaneous subarachnoid hemorrhage, artificial intelligence algorithm or not.

Where does this leave us? The study by Yang et al advances the field, given its large size and external validation. But the study seems to fall short of the ultimate goal of changing the way we interpret CT angiography in patients suspected of having aneurysms. More than a decade ago, some believed that CAD for mammography might even have deleterious effects on cancer detection in multiple ways (10). "Automation bias" from relying too much on the machine and "letting one's guard down" might also diminish sensitivity in aneurysm detection. Fatigue from looking at more than a dozen false-positive findings per case may suppress reader performance. Further, it is not clear if follow-up examinations in a given patient were excluded. If they were not, then the lack of a mention of this situation suggests that the same patient and aneurysm may have been present in both the training and validation sets.

As with all deep learning and CAD tools, there is a desperate need for large multi-institutional databases. For aneurysm, an external reference standard such as digital subtraction angiography would be ideal. Also, clinical data with subsequent rupture information could be useful to have artificial intelligence help predict risk of rupture in detected aneurysms. Understanding how best to integrate such tools into the practice is also needed. As with most CAD, an algorithm per se does not treat patients and, therefore, careful study of how to integrate the information into the practice is still needed.

Disclosures of Conflicts of Interest: **D.F.K.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed money to author's institution for grants from GE Healthcare; disclosed money to author for manuscript preparation from Superior Medical Experts; disclosed money to author's institution for patent for automated aneurysm detection entitled "Computer-aided detection (CAD) of intracranial aneurysms." Other relationships: disclosed no relevant relationships. **B.J.E.** disclosed no relevant relationships.

References

1. Arimura H, Li Q, Korogi Y, et al. Automated computerized scheme for detection of unruptured intracranial aneurysms in three-dimensional magnetic resonance angiography. Acad Radiol 2004;11(10):1093–1104.
2. Shi Z, Hu B, Schoepf UJ, et al. Artificial Intelligence in the Management of Intracranial Aneurysms: Current Status and Future Perspectives. AJNR Am J Neuroradiol 2020;41(3):373–379.
3. Stember JN, Chang P, Stember DM, et al. Convolutional Neural Networks for the Detection and Measurement of Cerebral Aneurysms on Magnetic Resonance Angiography. J Digit Imaging 2019;32(5):808–815.
4. Park A, Chute C, Rajpurkar P, et al. Deep Learning-Assisted Diagnosis of Cerebral Aneurysms Using the HeadXNet Model. JAMA Netw Open 2019;2(6):e195600.
5. Ueda D, Yamamoto A, Nishimori M, et al. Deep Learning for MR Angiography: Automated Detection of Cerebral Aneurysms. Radiology 2019;290(1):187–194.
6. Nakao T, Hanaoka S, Nomura Y, et al. Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography. J Magn Reson Imaging 2018;47(4):948–953.
7. Yang J, Xie M, Hu C, et al. Deep Learning for Detecting Cerebral Aneurysms with CT Angiography. Radiology 2021;298:155–163.
8. Kassell NF, Torner JC, Haley EC Jr, Jane JA, Adams HP, Kongable GL. The International Cooperative Study on the Timing of Aneurysm Surgery. Part 1: Overall management results. J Neurosurg 1990;73(1):18–36.
9. Wiebers DO, Whisnant JP, Huston J 3rd, et al. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. Lancet 2003;362(9378):103–110.
10. Philpotts LE. Can computer-aided detection be detrimental to mammographic interpretation? Radiology 2009;253(1):17–22.